CS/CS 316/365 Deep Learning

Activity 4

September 22, 2024

Loss Functions

Activity needs to be handwritten. Submission will be online on canvas only.

- Show that the logistic sigmoid function sig[z] becomes 0 as z $\rightarrow$ -∞, is 0.5 when z = 0, and becomes 1 when z $\rightarrow \infty$. Sigmoid's equation is given below. Showing doesn't require lengthy proof. Try putting in these values and show that result reaches to where it should be.

$$\mathrm{sig}[z] = \frac{1}{1 + \exp[-z]}$$

- The loss L for binary classification for a single training pair {x, y} is:

$$L = -(1-y)\log\Big[1 - \mathrm{sig}[\mathrm{f}[\mathbf{x}, \boldsymbol{\phi}]]\Big] - y\log\Big[\mathrm{sig}[\mathrm{f}[\mathbf{x}, \boldsymbol{\phi}]]\Big]$$

  where sig[z] is given above in first task. Plot this loss as a function of the transformed network output sig[f[x,ϕ]] $\in$ [0, 1] when the training label

  - y = 0
  - y = 1

- Consider a multivariate regression problem where we predict ten outputs, so y $\in \mathbb{R}^{10}$, and model each with an independent normal distribution where the means μ_d are predicted by the network, and variances σ^2 are constant. Write an expression for the likelihood Pr(y $\mid$ f[x,ϕ]). Show that minimizing the negative log-likelihood of this model is still equivalent to minimizing a sum of squared terms if we don't estimate the variance σ^2.